

**VINCULO ENTRE LOGRO EDUCACIONAL Y  
CARACTERÍSTICAS DE ESTÁNDAR DE VIDA EN LOS  
HOGARES DEL NORESTE DE ARGENTINA: UN ENFOQUE  
ECONOMETRICO CUALITATIVO**

CERNO, Leonel ([leonel.cerno@uem.es](mailto:leonel.cerno@uem.es))

Universidad Europea de Madrid

---

**Resumen**

En este artículo se procura estudiar el impacto que provoca una serie de características referidas a la familia y al ambiente en que ha crecido un habitante del noreste Argentino, sobre la probabilidad de alcanzar un cierto nivel educativo. Los resultados obtenidos indican que la educación, la condiciones familiares y el abandono de la escuela del jefe de familia tienen un impacto importante en el logro educativo del individuo de esta región de Argentina.

**Abstract**

In this article it is sought to study the impact that have about the probability empirically an individual of the Argentinean northeast to possess a certain formation level, a series of characteristics referred to the family and the environment in that this individual has remained. The obtained results indicate that the family boss's education, the conditions of the housing and the school abandonment have important impacts on the educational achievement of the individual of this region of Argentina.

*JEL classification:* C5, O54

Palabras Clave: Modelo logit, logro educacional, demanda de servicios sociales, escolarización.

---

**1. Introducción**

En sentido estricto la educación constituye una inversión en capital humano que a largo plazo proporciona un incremento de la productividad privada y social. Esta hipótesis fue sostenida inicialmente por Becker (1964) que derivaría luego en los llamados

*Schooling Models* de Mincer (1974) y tantos otros autores posteriores.

Para otros autores como Spence (1973) o Stiglitz (1975), la educación constituye una *señal* indicativa de la productividad innata del individuo, se realiza así un proceso de autoselección que culmina con el título universitario y esto permitirá agrupar a los individuos en el mercado de trabajo conforme a su capacidad productiva. Estos últimos se denominan modelos de *Screening*, y si bien parecen hipótesis contrapuestas con los más arriba mencionados, en la actualidad se puede admitir una coexistencia de estos procesos, según se ve en el trabajo de Corugero, García y Martínez (1992).

En este artículo se trata de reflejar y comentar algunos resultados de un estudio referido al grado de escolarización en los tres tipos de enseñanza (primaria, secundaria y universitaria) en el noreste argentino, concretamente en las provincias de Chaco, Corrientes, Formosa y Misiones. La forma de realizarlo es a través de contrastes empíricos de la probabilidad del logro alcanzado educacionalmente y una serie de características, todas ellas referidas al estándar de vida (capacidad de subsistencia, condiciones de la vivienda, servicios sanitarios y educación).

En un principio se incluyó además una variable continua referida a la población, pero se comprobó que esta no es influyente como para rechazar la hipótesis nula de parámetro igual a cero. El análisis se realiza a partir de datos del Documento de Trabajo N° 4 de la CEPA (Comité para el Estudio de la Pobreza en Argentina), realizado este a su vez de las bases de datos obtenidas del Censo Nacional de 1991.

El interés de ver porqué en esta zona de Argentina unas personas consiguen finalizar un determinado nivel de estudios y otras no, viene no solamente motivado por la importancia que tiene la educación en el bienestar económico de un adulto, sino además porque existe una relación de círculo vicioso de la pobreza según lo

planteaba Nurske (1953)<sup>1</sup> como se verá en las especificaciones del modelo planteado. Es necesario tener en cuenta que las cuantificaciones de estadísticos descriptivos de pobreza utilizan líneas imaginarias para poder definir a los que están del lado de los de menores ingresos de los que no lo están, todo esto a los fines prácticos de definición.

## **2. Marco Teórico**

En los estudios económicos existe una manera de explicar porqué unos individuos alcanzan un determinado nivel educativo y otros no a partir de modelos formales. Dichos modelos deben estar centrados en la teoría del comportamiento de la familia, ya que la familia es vista como una unidad de producción en la que los padres toman decisiones de inversión (en activos y capital humano de los hijos) y de consumo, todo esto ponderado según los recursos económicos de los cuales disponen. El logro educacional del individuo (el capital humano del mismo) necesariamente se va a ver afectado por la cantidad de recursos y su distribución en el tiempo.

De esta forma, algunos individuos tienen ventaja por el simple hecho de pertenecer a familias con recursos asequibles y otras características culturales y genéticas favorables.

---

<sup>1</sup> Afirma que en los países pobres, las propias fuerzas del mercado perpetúan la pobreza, y dado que para salir de ella se requiere invertir para aumentar la productividad, ello resulta difícil, no solo por el escaso ahorro de los pobres, sino por la falta de incentivos de beneficio para construir “plantas” de alta productividad cuando el mercado local existente para su producto es demasiado pequeño. El artículo seminal de esta idea pertenece a Rosenstein – Rodan sobre los países en desarrollo en donde argumentaban que las complementariedades y externalidades en la demanda y en la producción generaban la necesidad de una programación de las inversiones. Estos argumentos fueron ampliados ante la necesidad de un *Big Push* para que el proceso de desarrollo tuviera éxito. Se destaca además en aplicar el concepto de *desempleo disfrazado* descrito por neomarxistas como Joan Robinson, como problema persistente (más que cíclico) en los países en desarrollo.

### 3. Especificación Econométrica

#### *Modelo Logit Binario*

El modelo planteado en este artículo se basa en la maximización de la utilidad de unos padres que se preocupan por sus hijos, pero en un escenario de necesidad básicas insatisfechas. Se supone una aproximación al problema de los tres niveles de escolarización influidos por la *capacidad de subsistencia*, las *condiciones de la vivienda*, los *servicios sanitarios* y la *educación del entorno*. El modelo que se utilizará es un logit binario con datos agrupados cuya variable dependiente es si el individuo se encuentra en ese nivel de educación ( $y = 1$ ) o no ( $y = 0$ ). Este modelo se lo especifica teóricamente tal como sigue:

$$P_i = F(x_i^T \beta) = \frac{1}{1 + e^{-(x_i^T \beta)}} = \Lambda(x_i^T \beta)$$

en donde  $i = 1, 2, 3, \dots, N$ ,  $P_i$  constituye la probabilidad específica grupal de una de las dos alternativas consideradas (en este caso la probabilidad de que el individuo esté escolarizado en el nivel determinado),  $x_i$  es el vector de observaciones de variables explicativas deterministas y  $\beta$  es el vector de coeficientes.

Como todo modelo econométrico, necesitamos una perturbación aleatoria que provea aleatoriedad a la variable dependiente cualitativa. Para ello tenemos las probabilidades empíricas calculadas de la forma convencional tal como se ve a continuación:

$$\hat{P}_i = \frac{f_i}{N_i} = P_i + \varepsilon_i$$

Donde:

$\hat{P}_i$  : Probabilidad estimada.

$f_i$  : Casos favorables.

$N$  : Totalidad de casos.

$P_i$  : Probabilidad real.

$\varepsilon_i$  : Perturbación aleatoria.

A partir de estas dos ecuaciones se puede demostrar que<sup>2</sup>:

$$\log \left[ \frac{\hat{P}_i}{1 - \hat{P}_i} \right] = x_i^T \beta + u_i$$

$$u_i = \frac{\varepsilon_i}{P_i(1 - P_i)}$$

en donde:

$$Var(u_i) = \frac{1}{N_i P_i (1 - P_i)}$$

Esa varianza puede ser calculada utilizando la probabilidad estimada  $\hat{P}_i$  y corregida de heteroscedasticidad utilizando mínimos cuadrados generalizados. Este estimador es consistente y asintóticamente normal. A partir de este modelo teórico, el modelo a ajustar entonces será <sup>3</sup>:

---

<sup>2</sup> Mc. Fadden (1984).

<sup>3</sup> En este modelo no se incluye un término constante para no interpretar los coeficientes estimados como efectos diferenciadores de la media general. Se verá la significatividad de los coeficientes, el signo de los mismos y como todas las variables estarán medidas en las mismas unidades, observamos las diferencias relativas entre ellos.

$$\log \left[ \frac{\hat{P}_i}{1 - \hat{P}_i} \right] = \beta_0 X_1 + \beta_1 X_2 + \beta_2 X_3 + \varepsilon_{ij}$$

en donde los coeficientes no se interpretan como elasticidades, sino que reflejan el efecto del cambio en la variable en  $\log \left[ \frac{P_i}{1 - P_i} \right]$ , y la variación en probabilidad depende de la probabilidad original y de los valores iniciales de las variables independientes y de sus coeficientes. Específicamente:

$$\frac{\partial P_i}{\partial X_{ij}} = \frac{\beta_j \varepsilon^{-(x_i^T \beta)}}{\left[ 1 + \varepsilon^{-(x_i^T \beta)} \right]^2}$$

En el modelo que ajustaremos los parámetros corresponden a:

$\beta_0$  : Efecto relevante dentro de la probabilidad de estar escolarizado en determinado nivel con respecto a la capacidad de subsistencia.

$\beta_1$  : Efecto relevante dentro de la probabilidad de estar escolarizado en determinado nivel con respecto a la condición de la vivienda.

$\beta_3$  : Efecto relevante dentro de la probabilidad de estar escolarizado en determinado nivel con respecto al grado de educación del resto de la población.

### *Los Datos*

Como se dijo más arriba, las estimaciones son realizadas a partir de datos del Documento de Trabajo N° 4 del CEPA. Según este documento, la dinámica de los procesos de cambio económico, político y social que se han ido dando en las últimas décadas en la Argentina, ha generado un significativo aumento del requerimiento de indicadores que permitan cuantificar la demanda de servicios sociales.

El ámbito geográfico utilizado son las cuatro provincias nombradas más arriba, y donde cada una de estas está a su vez dividida en departamentos que en total corresponden a 75 en toda la región. De la recolección de datos se realiza una estratificación de acuerdo a las cantidades porcentuales de demanda de cada servicio social. Dichos servicios fueron también nombrados más arriba y recordemos que son “*capacidad de subsistencia*”, “*condiciones de la vivienda y servicios sanitarios*” y “*educación*”. A su vez se ajustará el modelo especificado teóricamente con tres variables dependientes distintas. Estas son las planteadas a continuación:

Cuadro 1: Variables dependientes Cualitativas

Variable	Concepto
$Y_2$	Tasa de escolarización entre los 6 y 12 años
$Y_3$	Tasa de escolarización entre los 13 y 17 años
$Y_4$	Tasa de escolarización entre los 18 y 22 años

Fuente: CEPA

La tasa de escolarización es el resultado del siguiente ratio:

$$T_i = \frac{[\text{gente con } i \text{ años escolarizada}]}{[\text{gente total con } i \text{ años}]}$$

Un supuesto a tener en cuenta es que en la población con esa edad no existe *sobre-edad* (cuando se tienen más años cumplidos que los correspondientes para el grado escolar que se cursa) ni *repitencia*.

Las variables deterministas a tener en cuenta se describen a continuación:

Cuadro 2: Variables Independientes

Variable	Concepto
$A_2$	Porcentaje de jefes de hogar con estudios primarios incompletos
$B_1$	Porcentaje de la población viviendo en viviendas deficitarias
$C_8$	Porcentaje de niños de 10 a 14 años que han abandonado la escuela

Fuente: CEPA

En cuanto a la definición de cada una de las variables, la  $A_2$  se escoge como un indicador de la capacidad de subsistencia y apunta a identificar hogares que, a partir de determinadas características de sus jefes, presenten una incapacidad potencial para la obtención de ingresos suficientes tanto para su sostenimiento como para invertir en la adquisición de capital humano en sus hijos.

Se considera que la modalidad de inserción en el mercado educacional y luego en el laboral está fuertemente ligada al nivel de escolarización alcanzado, esto hace que dichos niveles educativos bajos de los jefes de hogar se acompañen con inserciones ocupacionales en condiciones más desfavorables en términos de estabilidad y nivel de ingresos, lo que lleva a desfavorecer además la inversión en capital humano para sus hijos.

La variable  $B_1$  es escogida para referirse a la situación habitacional que expresan distintos grados de privación referidos al tipo de vivienda, el hacinamiento por cuarto, la procedencia y la provisión de agua, las características de los servicios sanitarios y el material predominante en los pisos. Se considera que esto influye sustancialmente en la escolarización pues no se dan las condiciones óptimas para educarse adecuadamente.

La otra variable a considerar es  $C_8$  la cual hace referencia al abandono escolar de niños que se encuentran en el intervalo de edad de 10 a 14 años. Esta variable es relevante sobre todo para la



escolarización superior, y esto se da así debido a que suele ocurrir en que las personas abandonen la escuela en la preadolescencia (justo al terminar la educación primaria) para dedicarse a tareas manuales. Indudablemente influirá en la escolarización si está presente esta situación en el entorno.

El tipo de muestreo realizado es el de conglomerados polietápicos con estratificación proporcional por nivel de hábitat. Los estratos están definidos y distribuidos tal como se ve a continuación:

Cuadro 3: Frecuencias relativas de los departamentos provinciales conforme a intervalos predefinidos de demanda de servicios sociales

Niveles	Estratos	Departamentos Provinciales por Estrato			
		Corrientes	Chaco	Formosa	Misiones
I	93.4% - 40.8%	28%	67%	78%	29%
II	40.6% - 29.1%	48%	29%	11%	47%
III	28.7% - 19.2%	20%	4%	11%	24%
IV	19.0% - 10.8%	4%	0%	0%	0%
V	10.7% - 2.6%	0%	0%	0%	0%

Vemos en el cuadro de arriba que la provincia de Corrientes es la que menos mal se encuentra en cuanto a la cantidad de departamentos con niveles altos de demanda de servicios sociales. Como contrapartida tenemos la provincia de Formosa en la cual prácticamente las  $\frac{3}{4}$  partes de sus departamentos poseen altos niveles de servicios sociales deficitarios. En cuanto a los valores que tomarán las variables, serán los siguientes:

Cuadro 4: Tipo y valores de las variables a utilizar

Escolarización. Variable ficticia ( $i = 1, 2, 3$ ) tal que:
$Y_i \begin{cases} 1 \text{ si toma niveles III, IV o V (niveles altos)} \\ 0 \text{ cualquier otra posibilidad} \end{cases}$
Capacidad de subsistencia. Variable ficticia tal que:
$A_2 \begin{cases} 1 \text{ si toma niveles III, IV o V (niveles bajos de necesidad)} \\ 0 \text{ cualquier otra posibilidad} \end{cases}$
Condiciones de la vivienda y servicios sanitarios ( $B_1$ ): Variable continua referida al porcentaje de población habitando viviendas con algún tipo de precariedad.
Educación. Variable ficticia tal que:
$C_8 \begin{cases} 1 \text{ si toma niveles III, IV o V (niveles bajos de abandono)} \\ 0 \text{ cualquier otra posibilidad} \end{cases}$

### *Resultados de Estimación y Comentarios Interpretativos*

En esta sección discutimos los resultados de los modelos ya planteados. La variable endógena es la transformación logística de la probabilidad empírica de estar escolarizado. Se utilizan tres variables exógenas para tres variables endógenas distintas, y se utiliza el criterio de razón de verosimilitudes para evaluar la significatividad conjunta de las variables exógenas <sup>4</sup>. He aquí los resultados:

---

<sup>4</sup> Esta es una diferencia importante con respecto al modelo clásico de regresión, ya que la bondad de ajuste no debe realizarse a través del coeficiente de determinación convencional  $R^2$ , pues presenta varios inconvenientes, por lo que se han propuesto un conjunto de medidas alternativas, entre ellas medir la significatividad conjunta con la razón de verosimilitudes.

Cuadro 5: Resultados de las estimaciones del modelo logit

Variable	$Y_2$	$Y_3$	$Y_4$
$A_2$	0,6320 (0,603)	2,7742 (2,435)	1,5934 (2,199)
$B_1$	-0,048 (4,170)	-0,060 (3,759)	-0,018 (3,566)
$C_8$	0,3501 (0,338)	0,3791 (0,376)	1,2812 (2,354)
$RV$	-16,96	-16,00	-39,58

Nota: Los paréntesis equivalen al estadístico  $t$

Para el caso de  $Y_2$  (escolarización primaria) solamente son significativas las condiciones de la vivienda y servicios sanitarios y además de forma inversa. Estas están representadas por la variable  $B_1$  (cuando más casas con algún tipo de precariedad existan, menos probabilidad de escolarización primaria motivada habrá), mientras que el hecho de que el jefe de familia no esté escolarizado, o el hecho de que haya abandono escolar en el entorno, son variables ficticias no significativas.

Esto tiene sentido y es importante matizar que esta variable independiente irá disminuyendo cada vez más a medida que va cambiando la variable dependiente. Será así porque, tal como se está viendo, a la hora de escolarizarse en niveles superiores de la escuela primaria entrarán a ejercer influencia el nivel de escolarización del jefe de familia (a niveles más bajos de cantidad de jefes de familia con primaria incompleta, mayor nivel de escolarización se logra) y el grado de abandono escolar del entorno (a niveles más bajos de abandono escolar, mayor probabilidad de lograr niveles de escolarización altos), tal como vemos para el caso de utilizar  $Y_3$  e  $Y_4$  como variables cualitativas dependientes.

Vemos por ejemplo para el caso de  $Y_3$  (escolarización en edad secundaria), además de la variable  $B_1$  también es significativa

la variable  $A_2$  que como se expresa más arriba se refiere al porcentaje de jefes de hogar con estudios primarios incompletos.

Vemos que a menos porcentaje de padres sin primaria completa, mayor es la probabilidad de estudios secundarios. Al ser estos coeficientes positivos, se comprueba lo que se sostiene más arriba de que la educación del padre tiene un efecto importante a partir de la escolarización secundaria en adelante, no dándose esta influencia en la escolarización primaria.

Para el caso de  $Y_4$  (escolarización a nivel superior), además de las dos variables explicativas ya mencionadas, también es influyente la variable referida al grado de abandono escolar en la época adolescente, es decir la variable explicativa  $C_8$ . Sólo en el último modelo es significativa esta variable, y esto tiene sentido ya que a la universidad o al terciario sólo irán aquellas personas que no hayan abandonado y hayan completado la escuela secundaria.

Si el análisis, en vez de enfocarlo a partir de las variables cualitativas utilizadas en forma endógena, lo realizamos además por las variables explicativas, vemos que para el caso de la educación del jefe de familia se aprecia, como ya se viene diciendo, un efecto positivo del nivel de estudios del padre sobre los hijos. Al pasar del individuo con padres poco escolarizados a otro cuyos padres hayan podido superar al menos la escuela primaria, la probabilidad de obtener un título secundario será mayor que la de poder escolarizarse aún más en un nivel superior casi en un 74%<sup>5</sup>.

Lo que resulta interesante de analizar son las comparaciones relativas de la variable  $B_1$ , en lo cuál se ve que en los tres modelos tienen una relación inversa en la probabilidad de escolarizarse, pero con diferencias importantes. Por ejemplo tenemos el caso de la escolarización secundaria en cuya probabilidad esta variable es en la

---

<sup>5</sup> Está referido a la diferencia relativa entre ambos coeficientes. Esta comparación es válida por tener ambos la misma unidad de medida.

que menos ejerce influencia en la probabilidad de escolarización, siendo la contrapartida la escolarización primaria en la que más ejerce, con una diferencia entre ambas del 20%. Esto quiere decir que es más difícil entrar en la escuela primaria a la edad correspondiente cuando se padecen necesidades en cuanto a las condiciones de vivienda y servicios sanitarios<sup>6</sup>.

Para el caso de la educación superior, esta variable es un 63% menos influyente, (no por ello no significativa), pero en la probabilidad de escolarización secundaria es en donde menos influye con una diferencia porcentual del 20% con respecto a la primaria y de un 70% con respecto al nivel superior. Esto significa que las condiciones de vivienda serán siempre importantes a la hora de decidir escolarizarse, pero su influencia será mucho mayor para el caso de la escuela primaria que en el resto de los niveles considerados.

Para el caso de la variable correspondiente al grado de abandono escolar en edad preadolescente,  $C_8$ , solo será significativa en el caso de determinar la probabilidad de escolarización superior (terciaria y/o universitaria). Esto es razonable ya que la cantidad de individuos que llega al nivel superior de estudios está *diezmada* por el abandono escolar, mientras que en el caso de la escolarización en primaria y en secundaria el abandono escolar *en la edad considerada* no es significativo.

Teniendo en cuenta esto último, ¿cómo entonces es posible que el contraste de razón de verosimilitudes ( $RV$ ), en los tres modelos nos diga que se rechaza la hipótesis nula de no significatividad conjunta?. Para reforzar la consideración de variables en los tres modelos ajustados, se procederán a elegir los modelos mejores aplicando el criterio de información de

---

<sup>6</sup> Se podría ampliar el estudio incluyendo variables que tengan en cuenta la sobre